Module Specification

Module Summary Information

1	**Module Title**	Introduction to Data Science
2	**Module Credits**	20
3	**Module Level**	5
4	**Module Code**	CMP5353

5	**Module Overview**

Data applicable to organisational bodies for decisions making process are increasing rapidly. Automation of tracking, the increase of new data types (e.g., social media, audio, and video), improved storage of electronic records, re-use of administrative records, and the outburst of modelling data have all amplified the availability of data. However, making full use of these data requires not only right storage and but also advanced analytical capabilities.
Data Science is a multidisciplinary field that deals with technologies, processes, and systems to extract knowledge and insight from data and supports reasoning and decision making under various sources of uncertainty.
The goal of Introduction to Data Science is to teach students how to answer questions with data. The module will enable students to conduct computational investigations needed for vital decisions and prediction of new trends and at the same time enhance students' employability skills.

Learning and Teaching

Students will learn the essential skills to manage and analyse data with case studies. They will learn concepts such as exploratory data analysis, statistical inference and modelling as well as linear algebra and linear regression. They will also learn the necessary skill to develop data products.

The module is designed as an interactive-taught lectures using a chain of group discussions/activities, quizzes as well as workshop/lab exercises using diverse statistical tools and techniques. All class material will be motivated with real life examples involving data. Students will learn and use R programming language as well as other programming language that could be used to analyse large data.

6	**Indicative Content**

This module is about learning from data, in order to gain useful and actionable predictions and insights. Pre-processing unstructured data and obtaining insights from the data through analysis presents many computational challenges.

The module will introduce students to the following:

- Data Science life cycle
- Data Wrangling
- Introduction to R
- Exploratory data analysis and visualization
- Linear Algebra
- Basic methods for modelling and prediction
- Statistical inference
- Data Science tools and applications

7	Module Learning Outcomes On successful completion of the module, students will be able to:	
	1	Analyse, interpret and effectively communicate Data Science lifecycle including processing different data type using statistical programming language/tools
	2	Learn, apply and professionally report data wrangling and visualisation techniques to clean and shape data for analysis
	3	Explore data using statistical methods to critically examine the validity of statistical analysis
	4	Demonstrate an understanding of basic regression model for the purpose of description and prediction

8	Module Assessment		
Learning Outcome			
	Coursework	Exam	In-Person
1, 2, 4	X		
3			X

9	Breakdown Learning and Teaching Activities
Learning Activities	Hours
Scheduled Learning (SL) includes lectures, practical classes and workshops, peer group learning, Graduate+, as specified in timetable	48
Directed Learning (DL) includes placements, work-based learning, external visits, on-line activity, Graduate+, peer learning, as directed on VLE	62
Private Study (PS) includes preparation for exams	90
Total Study Hours:	200